Policy Research Working Paper 5458

Eliciting Probabilistic Expectations with Visual Aids in Developing Countries

How Sensitive Are Answers to Variations in Elicitation Design?

Adeline Delavande
Xavier Giné
David McKenzie

The World Bank
Development Research Group
Finance and Private Sector Development Team
October 2010

Abstract

Eliciting subjective probability distributions in developing countries is often based on visual aids such as beans to represent probabilities and intervals on a sheet of paper to represent the support. The authors conducted an experiment in India that tested the sensitivity of elicited expectations to variations in three facets of the elicitation methodology: the number of beans, the design of the support (pre-determined or self-anchored), and the ordering of questions. The results show remarkable robustness to variations in elicitation design. Nevertheless, the added precision offered by using more beans and a larger number of intervals with a predetermined support improves accuracy.

This paper—a product of the Finance and Private Sector Development Team, Development Research Group—is part of a larger effort in the department to develop rigorous methodology to elicit subjective expectations in the field. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at xgine@worldbank.org and dmckenzie@worldbank.org.

Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How sensitive are answers to variations in elicitation design?[#]

Adeline Delavande, *Universidade Nova de Lisboa and RAND*

Xavier Giné, *World Bank and BREAD*

David McKenzie, *World Bank, BREAD and IZA*

Keywords: Subjective Expectations; Survey Design; Development
JEL codes: D84; C81; O12

[#] The research of Delavande was supported by a grant from the National Institute of Child Health and Human Development (R03HD058976).

1. Introduction

An increasing number of surveys in both developed and developing countries elicit the subjective expectations that individuals have about a wide range of future outcomes.[1] Such data are being used for understanding many economic and social behaviors, such as saving for retirement, contraceptive choice, migration decisions, and schooling.[2] In developed countries surveys usually collect subjective expectations by directly asking respondents questions like, "What is the percent chance your monthly income in one year will be below $1000?" (e.g. Dominitz and Manski, 1997). While this direct approach has been used in developing countries by McKenzie et al. (2007) and Attanasio and Kaufman (2009), in many settings it is felt that asking respondents with low levels of education directly for a probability is too abstract, and so visual aids are used to help individuals express probabilistic concepts. Such an approach could also be useful for surveying groups in developed countries with lower education levels, such as children or the poor.

The typical approach using visual aids is to give respondents physical objects like beans, balls or stones, and then ask them to allocate into intervals or bins in accordance with their subjective expectations of different events occurring.[3] In practice individuals are usually given 10 or 20 items to allocate, allowing them to express probabilities in units of 0.10 or 0.05. When eliciting a distribution, a major design issue is then how to specify a set of intervals for respondents to allocate these items to, and how to define the support over which expectations are elicited. Delavande et al. (2009)'s survey of the existing literature reveals two common methods for doing this. The first is to use a common, pre-determined support for all respondents. Pre-existing data or prior knowledge of the range of possible values of the outcome is used to define the support, and a relatively large number of intervals are often given within this support. The alternative method is to first ask individuals their perceived maximum and minimum for

[1] Manski (2004) and Hurd (2009) review the experience in developed countries, while Delavande, Giné, and McKenzie (2009) review the experience of developing countries.
[2] See for example Hurd et al. (2004) on retirement, Delavande (2008) on contraceptive choice, McKenzie et al. (2007) on migration, and Attanasio and Kaufman (2009) on schooling.
[3] Delavande and Rohwedder (2008) use a similar approach to collect expectation data in the U.S. over the Internet.

the outcome being studied, and then using these to define a relatively small number of self-anchored intervals within this range.

An additional feature of survey work in developing countries is that many of the surveys being undertaken are under the direct control of researchers, presenting researchers with considerable flexibility in how they implement expectations questions. But researchers must then decide how exactly to employ visual aids. The existing literature shows different researchers have made different choices in this regard (Delavande et al., 2009), yet there is no evidence as to how sensitive the subjective probabilities obtained are to these design choices. The purpose of this paper is to test how sensitive results are to these variations in design, and thereby guide future survey efforts in this area.

We carry out a methodological randomized experiment with boat owners in Tamil Nadu, India in order to test the sensitivity of expectations about future fish catches to three variations in elicitation design. First, respondents were randomly assigned to receive either 10 or 20 beans. Second, respondents are asked about the distribution of the value of future catches using both a pre-determined support with many intervals and a self-anchored support with only four intervals, with the order in which these were asked randomized. The advantage of the self-anchored support is that it asks respondents about the range of values which are relevant to them. The disadvantage however is that it requires real-time calculations by the interviewer which can be time-consuming and subject to interviewer calculation error so in practice the feasible number of intervals is limited. In contrast, a pre-determined support can accommodate more intervals but if the support is very heterogeneous across respondents, then intervals will be wide to encompass everyone's relevant range. Finally, when eliciting the self-anchored support, which is designed based on elicited minimum and maximum of the distribution of catches, we randomized whether respondents were first asked about the maximum or the minimum.

The target population provides an excellent setting for testing variations in expectations elicitation methodology. The survey respondents are boat owners, who are broadly representative, in terms of literacy and education levels, of the typical respondent in many developing country field surveys. In a survey conducted in July 2009, we asked expectations about the value of fish a boat owner expected to catch in a day in the month of August 2009. We can then use the realized distribution of daily catches during August to compare the elicited distribution to the realized distribution at an individual level. This is a unique feature of this setting which makes it particular suitable for testing methodology, since in many other cases one only observes one realization at the individual level (e.g. income in the month, whether they live or not to a given age) and not the distribution.

Our results should offer considerable comfort to researchers employing visual aids to elicit subjective expectations. We find respondents to be willing and able to answer questions using this format, and that almost all the answers received obey basic laws of probability. Similar distributions are obtained with 10 beans as with 20 beans, although we do find the distributions with 20 beans are more accurate, with respondents able to use more intervals of the pre-determined distribution to express their beliefs. The distributions elicited with a pre-determined support and many intervals are remarkably consistent with the self-anchored distribution with a small number of intervals, and the ordering of the minimum and maximum does not significantly change responses. We therefore conclude that the elicited distribution is robust to many of the key elicitation design decisions a researcher must make. Nevertheless, we do find the most accurate results are obtained using 20 beans and a pre-determined support, suggesting this design could serve as a default for future studies.

The remainder of the paper is structured as follows. Section 2 outlines the survey setting and the design of the methodological experiment. Section 3 then describes the basic results in terms of whether the results obtained obey basic laws of probability, and how the distributions elicited vary with differences in elicitation design. Section 4 then compares the accuracy of the different design choices to the realized distribution, and Section 5 concludes.

2. Setting and Survey Instrument

2.1. Sample and Setting

We carry out our methodological experiments using a sample of 272 boat owners from seven villages in the southern tip of Tamil Nadu, India. These boat owners were first surveyed in 2004 as part of an ongoing panel study by Giné and Klonner (2007). We chose this setting for the experiment for three main reasons. First, as we will explain further, these boat owners offer an excellent setting for comparing elicitation designs of subjective distributions, because the distributions can be compared to individual-specific distributions of realizations, rather than just point estimates. Secondly, the boatowners average 5.5 years of education, and so are broadly representative of the types of individuals with at most primary education for whom visual aids are most needed. A final reason was expediency – we could piggyback on the existing survey infrastructure and other data collected to carry out this experiment for low cost.

The seven villages which the boat owners come from each have a population of about 1,500, and most villages have neither a harbor nor a jetty, a fact that restricts operations to beach-landing boats only. All year-round operating vessels have a crew of one to four men and are operated by local households. All of these households belong to the exclusively catholic fishing community of the village, which used to belong to a particular fishermen's caste within the Hindu caste system before collectively converting about 400 years ago.

We ask expectations about the value of future catches, and so it is useful to understand the basic context in which fishing takes place and the main sources of uncertainty in the value of catches. Boat owners fish using either a traditional kattumaram, or a more modern fiber boat. The kattumaram is a raft-like vessel made of two Alphesia logs tied together with two crossbeams at the two ends. The boat owner typically goes fishing alone or accompanied by another household member or relative. The beach-landing fiber boat is a relatively new technology, with the first such boats appearing in Tamil Nadu in the mid-1990s. Fiber boats typically have a crew of four, some of which can be relatives of the boat owner and the rest are laborer-fishermen. Every crew member earns a daily minimum wage and a percentage of the value of catches.

On a typical day, boats leave the shore around 1 am and land back at the market place on the beach between 7 and 11 in the morning. There, local fish auctioneers and staff from the fishermen association market the catches to a group of buyers, which comprise local traders as well as agents of nation-wide operating fish-processing companies. Prices are fixed in advance every week depending on the season and variety of fish, so most of the variation in the value of catches on a day to day basis comes from quantity and variety rather than price. Fish that are too small under international legal minimum size standards are sold at the local market at a discount.

Our data come from three different sources. Boat owner characteristics for all 272 boat owners come from a household survey conducted in November 2007. We then fielded our methodological experiment in July 2009 to elicit expectations about August catches. Actual daily catches for August then come from the hand-written records kept by both the fishermen association and the auctioneers. Daily records for five individuals show that they did not go fishing in August, as they decided to work as laborers for large mechanized boats. In addition, we do not have daily catch data for two boat owners because the auctioneer they work for refused to share his records. Two individuals also refused to answer the expectations survey. As a result we have expectations data for 270 boat owners, and realizations of August catches daily for 263 boat owners.

Table 1 provides some basic demographic characteristics of the sample of boat owners. All boat owners are male, with an average age of 41 years. Mean years of schooling is 5.5, with 90 percent having 8 or fewer years of education. The average years of schooling for male adults 25 and over in the developing world in the year 2000 was

5.74 years, so our boat owners are broadly representative of the education levels prevailing in the developing world (Barro and Lee, 2000). The boat owners all started fishing as laborers and eventually became owners. On average, they have been boat owners for the past 14 years. Twenty-six percent of the sample own a katumarram, 66 percent a single fiber boat, and the remaining 8 percent own more than one fiber boat.

2.2. The Visual Aid

We designed a short module on subjective expectations which followed the typical format used in many developing country surveys: a visual aid was introduced to conceptualize probabilities, respondents were then asked several questions which measure whether they understand basic concepts of probabilities, and then expectations about a future outcome of interest (in this case fish catches) were asked. Within this basic design we vary three specific elements by randomizing respondents into 8 different groups (Table 2), which we describe in more detail below.

Respondents were asked to express their expectations about future events by picking out beans and placing them on a sheet of paper in accordance with their subjective probability of an event occurring. Specifically, the interviewer started out by explaining the concept of probability to the respondents using the introduction from Delavande and Kohler (2009) as follows:

I will ask you several questions about the chance or likelihood that certain events are going to happen. Here are 10 beans. I would like you to choose some beans out of these 10 beans and put them on the sheet of paper to express what you think the likelihood or chance is of a specific event happening. One bean represents one chance out of 10. If you do not put any beans on the sheet of paper, it means you are sure that the event will NOT happen. As you add beans, it means that you think the likelihood that the event happen increases. For example, if you put one or two beans, it means you think the event is not likely to happen but it is still possible. If you pick 5 beans, it means that it is just as likely it happens as it does not happen (fifty-fifty). If you pick 6 beans, it means the event is slightly more *likely to happen than not to happen. If you put 10 beans on the sheet of paper, it means you are sure the event will happen.*

A practical issue is how many beans respondents have been given. The most common choices in the existing literature have been 10 and 20 beans (Delavande et al, 2009), although exceptions exist, such as the use of 12 stones by Luseno et al. (2003). Ten and twenty beans are easily interpreted as probabilities, with the trade-off being between less respondent burden (both in terms of time and cognition) with ten beans against the potential for allowing more precision with twenty beans. The 10-bean format forces respondents to round their probabilistic beliefs to nearest 10 percent while the 20-bean format forces them to round to the nearest 5 percent. Manski and Molinari (2010) find that in the U.S. Health and Retirement Study, answers to expectations asked in percent chance format tend to be rounded to the nearest 5.

To evaluate how much difference the number of beans makes, half the sample of respondents (4 of the 8 treatment groups) were randomly chosen to receive 10 beans, and the other half 20 beans. For those given 20 beans, the introduction paragraph above was modified accordingly. After reading the introduction, the interviewer illustrated the concept further by using an example: they were told to suppose 5 black beans and 5 white beans were placed in a box, and then asked how likely it is that they would pick a black bean without looking.[4]

2.3. Questions to Measure Understanding of Probability

The expectations module then began with five simple questions to measure whether boat owners understand the concept of probability. The five questions were as follows:

1) Imagine I have 5 fishes, one of which is red and four of which are blue. If you pick one of these fishes without looking, how likely it is that you will pick the red fish?
2) How likely are you to go to (nearby town) sometime in the next two days?
3) How likely are you to go to (nearby town) sometime in the next two weeks?

[4] Again this was modified for those respondents given 20 beans.

4) How likely do you think it is that you will not catch any fish in the month of August if you go fishing 6 days a week?
5) How likely it is that you will eat fish at least once during the month of August?

The first question serves to assess numeracy, and whether respondents can express a known probability with the visual aid. The second and third questions can be used together to evaluate whether respondents respect a basic property of probability by asking about nested events: if respondents understand the concept of probability, they should allocate a larger number of beans (or at least as many) to the likelihood of going to town in the next 2 weeks as they do to the next two days. Finally questions 4 and 5 ask about events which are likely to be zero probability (catching no fish in the month) and certain events (eating fish at least once in the month) for the target population, allowing us to see whether they can use the visual aid to accurately represent zero and unity probability events.

2.4. Eliciting the Distribution

After these preliminary questions, respondents were asked to report their subjective distribution of future catches during one day in August by allocating the total number of beans on a sheet of paper divided into intervals to express the likelihood that the catches will fall into various intervals, expressed in Rupees. We compare two different methods for providing the support and number of intervals that should be used for this: a pre-determined support with a large number of intervals, and a self-anchored support with a small number of intervals. These reflect the main two approaches used in the literature.

The pre-determined support is the same for all respondents, and contained 20 intervals. These intervals were chosen based on previous catch data, which gave the range of reasonable possibilities for the value of daily catches.[5] The first interval ranged from 0 to 150 rupees, with the first 14 intervals (up to 2100 rupees) all having width of

[5] When prior data is not available, one can use pilot testing to obtain a reasonable range.

150 rupees. The intervals then became increasingly wider to cover a range of extreme outcomes, with the last interval open-ended (5001 or more rupees). The self-anchored support is individual-specific. Respondents were first asked the maximum and the minimum value of catches that they would expect in a single day in the month of August. These responses are then used to construct four intervals of equal size. The left outer interval starts at the elicited minimum and the right outer interval ends at the elicited maximum.

The basic trade-off between the two methods is that while the self-reported support asks respondents about a range of values which are relevant to them, the disadvantage is that it requires real-time calculations by the interviewer when a paper survey is used, which can be time-consuming and subject to interviewer calculation error.[6] Moreover, since the bounds of the intervals are constructed based on the elicited maximum and minimum, they often are less rounded than those which one would use in a pre-determined support. For example, if the minimum and maximum answers are 100 and 2750, the intervals would be [100-762.5), [762.5-1425), [1425-2087.5), and [2087.5-2750]. Such intervals are likely to be harder for respondents to think about than intervals which are divisible by 10, 50 or 100.[7] This is also why the number of intervals has to be limited, since taking midpoints within these intervals requires more calculations and results in intervals which may not be easy for respondents to think about. In contrast, a pre-determined support enables the use of more intervals without increasing survey time or the risk that an interviewer makes a mistake. However, if the support is very heterogeneous across respondents, then a pre-determined support may require relatively wide intervals in order to cover everyone without having to employ too many separate

[6] A few surveys in developing countries now use PDAs or ultra-mobile computers for data collection. In principle this offers the possibility of facilitating more intervals with a self-anchored design through automatic calculations and rounding. However, the only application we are aware of to do this is ongoing work by Fafchamps et al. (2010) which asks self-employed individuals in Ghana their expectations over future sales, using six self-anchored bins rather than the more common four.

[7] Of course in practice there is no reason one could not have an additional step of rounding or mapping from these self-anchored intervals to intervals in a range which is divisible by 10 or 50. This has been done in some cases (e.g. Dominitz and Manski, 1997 and McKenzie et al. 2007), but it can further increase the interviewer burden.

intervals, limiting the usefulness of the elicited distribution (see also the discussion in Delavande et al., 2009).

In order to be able to compare the distribution for the same individual under the two different methods, each individual was asked to provide their subjective distribution of the value of future catches using both methods. Since ordering and anchoring may influence how respondents answer the questions, the order of the elicitation of the two distributions was randomized: half of the respondents were first asked about the pre-determined support distribution, while the other half was first asked about the self-anchored individual-specific distribution. We can then compare the subjective distributions for these two halves of the sample, as well as compare the distributions under the two different methods at the individual level.

Finally, there are many examples in the social sciences where survey responses have been found to be sensitive to anchoring (e.g. Tversky and Kahneman, 1974). One case where this may arise in subjective expectations design is in asking the maximum and minimum, which are used to generate the intervals for the self-anchored distribution. It might be possible, for example, that asking about the maximum first anchors individuals into thinking about a set of good outcomes, while asking about the minimum first anchors them into thinking about bad states of the world. We therefore randomize the ordering of whether the maximum or minimum is asked first, as well as whether they are first asked to answer on the pre-determined support before giving either a maximum or a minimum. In total this then gives 8 different treatment groups, according to whether they receive 10 or 20 beans, whether they get the pre-determined support or self-anchored support question first, and whether they are asked the maximum first or the minimum. Although there are only 34 respondents in each treatment group, the treatment assignments are cross-cutting. We can therefore look at the impacts of individual design elements by comparing one half of the sample to the other.

3. Results

Respondents were willing to express their beliefs using the bean format, with very low item non-response. Out of the 272 respondents who were surveyed, two refused to answer any probability questions. All other 270 respondents answered the first five preliminary questions. Out of these 270 respondents, three did not answer any of the questions eliciting their distributions, or the minimum and maximum. One respondent was mistakenly given 14 beans rather than 10. When describing the subjective distributions, we exclude these four respondents.

3.1. Do Boat Owners Understand Basic Properties of Probabilities?

Before comparing the distributions elicited under different designs, it is useful to begin by checking that the respondents understand the concept of probability. The five questions outlined in Section 2.3 can be used for this purpose. The results show high comprehension of the concept of probability and the ability of respondents to use the visual aid of beans to express simple probabilities accurately. All but one respondent answered correctly a probability of one-fifth (i.e., 2 beans in the 10-bean format or 4 beans in the 20-bean format) to the red and blue fish question. In addition, all respondents allocated zero beans when asked about the zero-probability event, and all the available beans when asked about the certain event.

Figure 1 then plots the elicited probabilities of going to town within two days and within two weeks. The figure reveals several interesting aspects of respondents' answers. First, the range of answers shows respondents' willingness to use the whole scale from zero to one. Moreover, when respondents were given 20 beans and so had the opportunity to give an answer ending in 0.05, many did: 31% (17%) of the 20 bean treatment groups used an answer ending in 0.05 for the two-day (two-week) question respectively. Second, all of the points lie above the 45 degree line, showing that respondents obey the nesting property in their answers, giving a higher probability of going to town in two weeks than in two days.

Third, we do not find the high levels of heaping at focal points that have occurred in some of the previous literature. Expectations questions are often found to exhibit heaping at focal answers of 0%, 50% and 100% (e.g. Hurd and McGarry 1995), and responses of 50% have been shown to reflect uncertainty (Bruine de Bruin et al. 2000). In contrast, Figure 1 shows a coherent pattern of answers. No respondent allocated a probability of zero to the likelihood of going to town in the near future, reflecting that boat owners go there frequently. The most common answer, provided by 20% of the respondents, for the likelihood of going within the next two days is 0.5. However, less than 1% of the respondent answered 0.5 when asked about the likelihood of going to town in the next two weeks. The difference of rate of answers at 0.5 between the two questions suggests that respondents who answered 0.5 most likely meant a probability equal to one half, rather than expressed mere uncertainty. Otherwise, we would have probably seen a heaping at 0.5 for the likelihood of going to town in the next two weeks as well. For the two-week period, the most common answer is 1, provided by 53% of the respondents. However, given how common it is for boat owners to go into town during this period, this also seems like a genuine response rather than just a focal answer.

Overall, the answers to the 5 preliminary questions show that the boat owners are able to use the bean format to express the correct probability when asked about simple events, respect the monotonicity of nested events, use the whole range from zero to 1, and do not excessively use focal answers. As such, they are therefore a useful sample for exploring the consequences of differences in elicitation design.

3.2. Ten Beans Versus 20 Beans

Table 3 compares the subjective probabilities elicited using 10 beans to those using 20 beans. Panel A begins by comparing the distribution of responses to the likelihood of going to town in the next two days for the 10 bean treatment groups to that of the 20 bean treatment groups.[8] The distributions are quite similar, especially in central tendency. The medians are exactly the same, and we cannot reject equality of the means

[8] Results are similar for the two week formulation, so for brevity we discuss only the two day results here.

(p=0.671). However, the spread of the distribution is slightly smaller under the 20 bean format. For example, the interquantile range is 0.30 in the 10-bean design compared to 0.15 in the 20 bean design, and the standard deviation is 0.17 in the 10-bean design compared to 0.14 in the 20-bean design. One interpretation of this is that respondents have relatively concentrated beliefs around the mean of the distribution, but that, having fewer beans may constrain their answers. For example, while 12% of the respondents answered a probability of 0.3 in the 10-bean format, 8% allocated a probability of 0.3 and 5% allocated 0.35 in the 20-bean format.

Panel B of Table 3 then compares the distributions of responses that individuals give to their subjective expectations about future catches when using the self-anchored support. Recall in this format that there are four individual-specific intervals that respondents must allocate beans to in accordance with their beliefs as to the likelihood of future catches falling in each range. In general we again find quite similar distributions using 10 beans as 20 beans. For example, the first rows of the table show that the mean probability allocated to the first interval (from the minimum to halfway between the minimum and the midpoint) was 0.34 for the 10 bean treatments, and 0.36 for the 20 bean treatments, and that we cannot reject equality of means (p=0.188). However, many of the respondents who were given 20 beans used these to express their beliefs in a more refined way. For example, 52% of these respondents used an answer with an increment of 0.05 when asked to provide the probability that future catches will fall in the first interval. As a result, we can reject equality of distributions for some of the intervals, particularly the top interval (from halfway between the midpoint and the maximum to the maximum). Nonetheless, the distributions are still quite close overall, and they do suggest that individuals are not experiencing more problems using 20 beans to express probabilities than with 10 beans.

Finally, Panel C of Table 3 then compares the distributions of responses that individuals give about future catches when using the pre-determined support and larger number of intervals. We do find that respondents allocate positive probability mass to more intervals when they are given more beans: the mean (median) number of intervals used is 13.8 (14) in the 20-bean format, compared to 9.3 (9) in the 10-bean format. The 5th percentile in the 10 bean format is 8 intervals, and the 25th 9 intervals. Thus with 10 beans, respondents are almost providing a uniform distribution over a subset of intervals, whereas with 20 beans they are placing more mass in some intervals than others. Thus in practice when using fewer beans, one should consider using fewer intervals – which means using wider intervals and obtaining less precision.

Table 3, panel C combines the 20 intervals into four groups for ease of display: the tails, which are the first two intervals (<300 rupees), and the last two intervals (>4000 rupees); and then intervals 3 to 10 (from 301 to 1500 rupees) and intervals 11 to 18 (1501 to 4000 rupees). We see the allocation of mass is quite similar in the middle intervals, and cannot reject equality of mean, medians, or distributions for these middle intervals. In contrast, for the tails we do reject equality of medians, and for the first two intervals, equality of distributions. Respondents are not allocating much mass to the tails, but when they do, often allocate 0.05 when they have 20 beans. Thus with more intervals, there is a difference between using 20 beans and 10 beans.

3.3. Pre-determined Support and Many Intervals Versus Self-anchored Support with Few Intervals

Next we compare the subjective distributions obtained using the pre-determined support which is the same for all respondents to those with the self-anchored support, which are individual-specific. We can compare the results at the treatment group level, by comparing the distribution obtained for the half of the sample which were asked the pre-determined support first to the distribution for the half the sample which were asked the self-anchored support first. This enables us to see whether, on average, we get the same distributions regardless of method. We can then examine consistency of distributions at the individual level, enabling us to see whether for each individual the distribution elicited looks the same regardless of the method used for eliciting it.

We begin by comparing the distributions at the treatment group level. It is not clear how best to compare distributions of distributions, especially because the self-anchored distribution has individual-specific intervals. We use two approaches. First, in panel A of Table 4, we calculate the mean and median value of catches from the elicited distributions. To do this we assume a continuous uniform distribution within each interval.[9] We can then compare whether we get similar means and medians under the two different elicitation methods. We see the means are extremely similar using the two approaches: the mean of the subjective means is 1,910 rupees with the self-anchored support compared with 1,909 rupees with the pre-determined support (p=0.991). The subjective medians do differ – the mean subjective median is 1,750 rupees with the self-anchored support, compared to 1,642 rupees with the pre-determined support. However, this difference likely arises in large part from the parametric assumption (a uniform distribution within intervals) that yields the self-anchored support distribution based on four intervals to be of a rather different shape than the pre-determined support distribution that has 20 intervals.

A second approach to comparing distributions is to re-classify the 20 common intervals of the pre-determined support into 4 individual-specific intervals. We define these to match as closely as possible the individual-specific supports that a respondent has for their self-anchored questions. For example, an individual whose intervals on the self-anchored distribution are [200, 1400), [1400, 2600), [2600, 3800), [3800, 5000] has the 20 intervals in the pre-determined support collapsed and reclassified into the four following intervals: [0, 1350], [1351, 2500], [2501, 4000], [4001,∞). Panel B then compares the amount of probability mass assigned to these similar intervals. The distributions in these intervals are very similar, and we cannot reject equality of means, medians, or distributions.

We can also investigate more directly whether the distributions elicited using the two methods are coherent with one another at the individual level. Denote the intervals of

[9] So the mean is then the sum of the probability mass assigned to each interval times the midpoint of that interval. The last interval of the pre-determined support distribution is open-ended (5001 or more rupees). To compute the mean, we assume that it is bounded by 6500 rupees.

the individual-specific support of a given respondent by I_1, I_2, I_3 and I_4. We then combine intervals from the pre-determined support to construct new intervals C_i (i=1,2,3,4), defined to be the smallest intervals which contain each I_i. For example, if I_1=[200, 650], then C_1=[151, 750]. If the two elicited distributions are consistent with one another, then $Pr(I_i) \leq Pr(C_i)$ for i=1 to 4. Remarkably, only 10 respondents (3.8%) provided distributions inconsistent with each other. Note that there was only one set of beans, so the respondents could not see their answers to the first distribution when answering questions about the second distribution.

By eliciting the same distribution using two methods we can also assess whether the elicited minimum and maximum can be interpreted as truly lower and upper bounds of the subjective distribution by checking whether respondents allocate beans to pre-determined intervals that fall entirely below the self-reported minimum and entirely above the self-reported maximum.[10] Among the respondents for which we can evaluate this, only one respondent allocated beans below the minimum, and 6 respondents allocated some beans to intervals above the maximum. In addition, the total beans allocated in these intervals translate into probability mass that never exceeds 10 percent. These numbers are smaller than those reported in Dominitz and Manski (1997) and Delavande et al (2009) which look at this issue in the context of eliciting the distribution of yearly income in the U.S. and Tonga respectively. Perhaps boat owners are more used to thinking about daily catches than respondents in these studies are about yearly income.

Overall these results show that the distribution is very similar when using a pre-determined support to using a self-anchored support. Note however that these comparisons are essentially collapsing the richer information in the pre-determined support down to intervals and moments that can be compared with the self-anchored distribution. As such, a direct comparison of the two does not enable us to see whether the finer degree of detail in the pre-determined support is useful. This question is

[10] For 21% (29%) of the respondents, the elicited minimum (maximum) falls in the first (last) interval. For those respondents, we cannot evaluate whether they provide probability mass below and above the elicited minimum and maximum.

addressed in Section 4, when we compare the accuracy of the two methods of providing a support and intervals.

3.4. How Does the Ordering of Questions Affect Responses?

Next we examine how sensitive the answers to questions on the maximum and minimum are to ordering, which might influence these responses through anchoring. First, we can compare how the answers vary with whether the minimum or the maximum is asked first. We find the ordering has no impact on the answers. The mean (median) subjective minimum is 233.5 (200) for respondents who were asked the maximum first, and 236.6 (200) for respondents who were asked the minimum first. Similarly, the mean (median) subjective maximum is 4517 (4500) for respondents who were asked the maximum first, and 4599 (4600) for respondents who were asked the minimum first. We cannot reject equality of mean subjective minima ($p=0.766$) or equality of mean subjective maxima ($p=0.587$). Likewise, neither can we reject equality of medians, or equality of distributions of these subjective minima and maxima.

Second, we can see whether asking individuals to specify their future catch distribution on the pre-determined support first affects the answers they give to the maximum and minimum, relative to them being asked these maximum and minimum without first thinking through the full distribution. Table 5 compares the distributions of the subjective maximum and minimum across these two orderings. The distributions are again seen to be strikingly similar. For example, the mean minimum is 234.3 when these questions are asked first, and 235.7 when the minimum and maximum are asked after the pre-determined support questions. We cannot reject equality of means or of distributions. Thus it does not appear that answers to the pre-determined support question are anchoring how individuals respond to the minimum and maximum.

In addition to examining the sensitivity of the elicited maximum and minimum to ordering, we can also test whether asking the pre-determined support distribution before the individual-specific support influences individuals' answers for each of these distributions. Again, we find no impact of the ordering of the questions on the patterns of answers. The mean and percentiles of the distribution of probability mass in each of the four intervals of the individual-specific distribution for each ordering are nearly identical. The same is true for the pre-determined support distributions.

3.5. What Do We Conclude about the Sensitivity of Subjective Expectations to Design?

Overall the results of this Section suggest that the subjective distribution elicited is quite robust to modifications in the elicitation design. In particular, the pre-determined support and self-anchored support give strikingly similar subjective distributions at the individual level, and the ordering of questions does not significantly affect responses. We do find similar responses from 10 beans as for 20 beans with simple probability point estimates (the probability of going to town in 2 days), and using the self-anchored support with 4 intervals. However, when using a pre-determined support with many intervals, we find that individuals with 20 beans use more intervals, and have a distribution which is less uniform. It seems that allowing 20 beans instead of 10 beans does not seem to unduly increase respondent burden, and allows more precision and nuance in expressing beliefs. In the next section we then ask whether this results in more accurate subjective distributions.

4. Do Boat Owners Have Accurate Expectations about Future Income?

We have seen that the different elicitation methods do give quite similar subjective distributions, but that there are some differences. The question then arises as to which method is best capturing true subjective expectations. Delavande et al. (2009) outline several categories of checks which are often used in the literature to ascertain whether the subjective expectations being measured "work". We have seen both methods satisfy basic criteria such as having high response rates and being internally consistent. The two main criteria that are commonly used for assessment are then accuracy, and ability to predict choice behavior. We believe both are important, but in the current setting that accuracy is the better measure for comparisons. One reason for this is that we do not have good measures of choice behavior to look at – the boat owners all fish each day for a similar number of hours, so we would not expect to see differences in fishing behavior. One might imagine differences in consumption or savings behavior could arise from differences in expectations of future income, but we do not have good measurements of these, and the sample size is likely too small to detect effects in any event. In contrast, the data here are very rich for assessing accuracy, since they contain an individual specific distribution of realizations of catches. As a result, we can compare a respondent's subjective distribution of future catches for one day in August to the distribution of realized catches for the whole month of August.

4.1. Descriptive Comparisons of Actual Realizations to Subjective Distributions

Table 6 compares the distribution of responses that individuals give about future catches when using the pre-determined support to the individual distributions of realizations. As in Table 3, we collapse the 20 intervals into four groups for presentation purposes. Overall, the percentiles of the distributions of subjective and realized distributions look very similar, indicating that boat owners have reasonably accurate expectations on average.[11] If we look at the means, we find that individuals tend to allocate slightly more probability mass to the larger intervals compared to the realizations, suggesting that some individuals were overly optimistic about their future catches.

Figures 2 and 3 provide a second form of visual comparison between the subjective expectations and the realizations. Figure 2 plots the mean of the subjective self-anchored support distribution (derived under the assumption of a uniform distribution within intervals) against the actual mean of catches for the month of August for each respondent.[12] The dots are concentrated around the 45 degree line for respondents with subjective mean below 2000 Rupees, but they are more dispersed and concentrated slightly above the 45 degree line at higher values of the subjective mean,

[11] There were no important weather shocks or other such events during this period that would lead one to worry about aggregate shocks in comparing realizations to expectations.
[12] The graph excludes one respondent whose actual mean was above 5,000.

suggesting again that some individuals were excessively optimistic about large catch values.

The pre-determined and self-anchored supports show similar accuracy in terms of matching the means of the empirical distributions. In contrast, Figure 3 shows that the pre-determined distribution does a better job of matching the standard deviation of the realized distribution, with the points quite tightly clustered around the 45 degree line. The standard deviations estimated using the self-anchored distribution typically overstate the true standard deviation. This difference shows the advantage of asking respondents to provide more precise information by giving them more intervals. It also provides evidence that the standard deviation of the elicited distribution is a good proxy for the standard deviation of the actual income process, which cannot be computed without a sufficiently long time series.[13]

4.2. Which Designs Are More Accurate?

We now turn to formally comparing the accuracy of the different designs. The measure of accuracy that we use is based on the absolute value of the area between the cumulative distribution function of actual and expected catches in August. This statistic is similar to the Kolmogorov-Smirnov measure defined as the maximum value of the absolute difference between both cumulative distribution functions.[14] Since the last interval in the pre-determined support is open ended, an assumption must be made about its range when computing the statistic. We assume that the upper bound is 6,500Rs. Since relatively little mass is allocated to this last interval, the results are robust to other reasonable assumptions about how to value this interval.

[13] Attanasio (2009) uses the standard deviation of the subjective distribution as a proxy for the standard deviation of household income. Our finding provides some rational for doing so.

[14]We favored this measured over the chi-squared statistic for binned data (Press, 1992 and Gine et al, 2009) for two reasons. First, because catches are continuous and we prefer to avoid unnecessary binning and second, because the chi-square statistic is very sensitive to the number of intervals used, which is key variable that distinguishes both designs. .

Table 7 reports the results from regressing the logarithm of our statistic measuring accuracy against a set of experimental dummies and boat owner controls. The experimental dummies indicate the various randomized treatments used for eliciting distributions. For example "First distribution elicited" takes value one for observation *i* if the support used in calculating the statistic for observation *i* was the first support over which expectations were elicited. The dummy variable "Maximum was elicited first" takes the value 1 if the maximum of the support was asked first, and zero otherwise. Thus, it takes value zero when the distribution with a pre-determined support was elicited. Columns 1-4 pool the data from both the pre-determined and self-anchored supports, columns 5-7 report the results for the pre-determined support and columns 8-10 report the results for the self-anchored distribution. Columns 1, 5 and 8 run a simple regression using only the experimental dummies as regressors. Columns 2, 6 and 9 add village fixed effects while columns 3, 7 and 10 include boat owner characteristics as well as village fixed effects. Because the design variations were randomized, including these additional controls should increase precision, but have little effect on the coefficients on the design treatments. Column 4 also includes the interaction of "support is pre-determined" and "Number of beans is 20" because allowing a larger number of beans should have a more pronounced effect when the number of the number of intervals is large. In columns 1-4, standard errors are clustered at the boat owner level because there are two observations (one for each distribution elicited) per boat owner.

Note that the larger the dependent variable the greater the discrepancy between the expected and realized distribution. Thus a positive and significant coefficient indicates that the variable decreases accuracy, while a negative coefficient indicates that the variable improves accuracy.

The results show that using a pre-determined support rather than the self-anchored support results in a 21-25 percent improvement in accuracy. Given the similarity of distributions when collapsed to the same number of intervals, we interpret this as showing that allowing for more intervals does improve accuracy. We also see that using 20 beans rather than 10 beans improves accuracy, especially when combined with the pre-determined support. Thus greatest accuracy occurs when using 20 beans and the pre-determined support, which results in a 38 percent improvement in accuracy relative to using 10 beans and a self-anchored support. Allowing respondents the opportunity to provide more precise answers does therefore seem to have significant value in terms of accuracy.

We do find weak evidence of some difference in accuracy from ordering. When the minimum is elicited first, respondents tend to be 7 to 8 percent less accurate than when they are asked the maximum first. However, this difference is only, at best, significant at the 10 percent level. In terms of boat owner characteristics, education matters in a non-linear (U-shape) form. That is, more educated individuals have more accurate expectations, with the accuracy gains from education exhibiting decreasing returns.[15] We see no significant effect of age, type of boat, or the number of boats owned in the accuracy of expectations.

5. Conclusion

Researchers attempting to measure subjective expectations with populations for whom probabilities and percentages are not completely familiar concepts have relied on visual aids to elicit subjective probabilities and distributions. They face a number of elicitation design decisions when doing this, and currently it is not clear how robust the expectations elicited are to the range of variations in design typically seen in practice. This paper has reported on several experiments which were conducted with the aim of testing the sensitivity of subjective expectations to these design choices. The results show considerable robustness of the responses to variations in the number of beans respondents are given to express probabilities, to whether they are given a pre-determined support with many intervals or a self-anchored support with few intervals, and to the ordering of questions used to anchor respondents. This should provide researchers with additional confidence in the data collected from such questions.

[15] The minimum of the quadratic function for education is estimated at 6.6 years, so the majority of boat owners (70 percent) have education levels left of the minimum.

Nevertheless, we do find that accuracy is greatest when 20 beans instead of 10 are used together with a pre-determined support with many intervals. More beans and more intervals allow respondents the opportunity to be more precise in their responses, and we do find this to occur in practice. Respondent burden did not seem markedly greater using this approach to other methods, suggesting that researchers should consider allowing respondents the opportunity to be more precise in their responses in other endeavors to measure expectations.

Of course for practical reasons one cannot use an arbitrarily large number of beans and intervals. And it is still unclear at this point whether using more than 20 beans would be implementable and useful. Using a visual aid based on less than 100 beans inevitably forces respondent to round their answers more than when using a percent-chance wording, which has been the standard approach in developed countries. However, the visualization of the distribution provided by the bean format may help respondents and conceivably lead to less rounding than would occur with a percent-chance wording. Which elicitation methods, in both developed and developing countries, lead to less rounding is an interesting question for future research.

A caveat is that these results come from one setting in one country. The question is then how generalizable such findings may be. This can only be truly answered by further studies like this one in different settings. However, the boat owners studied here have 5.5 years of education on average, which is a similar level to many poor individuals in developing countries for whom these visual aids are intended. It is thus perhaps not unreasonable to believe that the results may be broadly indicative of the ability of people with these education levels to answer such expectations questions. A second point to note, however, is that our survey asked expectations about commonplace events, which individuals had good knowledge of. There is evidence from other settings (see Delavande et al, 2009) that subjective expectations are decidedly less accurate when individuals are asked about events which they have not yet had much experience over. It is possible that variations in elicitation design may matter differently in such cases, and we see testing this as a fruitful area for future methodological experiments.

References

Attanasio, Orazio (2009) The Empirical Implications of Self-Enforceable Contracts; Measuring Sticks and Carrots in Rural México, mimeo UCL. .

Attanasio, Orazio and Katja Kaufmann (2009) "Educational choices, subjective expectations, and credit constraints", *NBER Working Paper No. 15087.*

Barro, Robert and Jong-Wha Lee (2000) "International Data on Educational Attainment: Updates and Implications", CID Working Paper no. 42.

Bruine de Bruin, Wandi, Baruch Fischhoff, Susan G. Millstein and Bonnie L. Halpern-Felsher, "Verbal and Numerical Expressions of Probability. 'It's a Fifty-Fifty Chance'," *Organizational Behavior and Human Decision Processes* 81 (2000), 115-31.

Delavande, Adeline (2008) "Pill, Patch or Shot? Subjective Expectations and Birth Control Choices", *International Economic Review* 49(3): 999-1042.

Delavande, Adeline, Xavier Giné and David McKenzie (2009) "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence", Mimeo. World Bank.

Delavande, Adeline and Hans-Peter Kohler (2009) "Subjective Expectations in the Context of HIV/AIDS in Malawi", *Demographic Research* 20(31): 817-74.

Delavande, Adeline and Susann Rohwedder, (2008). "Eliciting Subjective Expectations in Internet Surveys," *Public Opinion Quarterly*, 72(5): 866–891.

Dominitz, Jeff and Charles Manski (1997) "Using Expectations Data to Study Subjective Income Expectations", *Journal of the American Statistical Association* 92(439): 855-67.

Fafchamps, Marcel, David McKenzie, Simon Quinn and Christopher Woodruff (2010) "How much can using PDAs increase precision of profits and sales measurement in panels?" Mimeo.World Bank

Giné, Xavier and Stefan Klonner (2007) "Technology Adoption with Uncertain Profits: The Case of Fibre Boats in South India", Mimeo World Bank.

Giné, Xavier, Robert Townsend and James Vickery (2009). "Forecasting When it Matters: Evidence from Semi-Arid India" Mimeo World Bank.

Hurd, Michael. (2009), "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1, 543-564.

Hurd Michael and Kathleen McGarry. (1995). ""Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." *Journal of Human Resources* 30:268–92.

Hurd, Michael, James P. Smith and Julie Zissimopoulos (2004) "The Effects of Subjective Survival on Retirement and Social Security Claiming", *Journal of Applied Econometrics* 19(6): 761-75.

Luseno, Winnie K., John G. McPeak, Christopher B. Barrett, Getachew Gebru and Peter D. Little (2003), "The Value of Climate Forecast Information for Pastoralists: Evidence from Southern Ethiopia and Northern Kenya," *World Development*, vol. 31, no. 9, pp. 1477-1494

Manski, Charles (2004) "Measuring Expectations," *Econometrica* 72(5): 1329-76.

Manski, Charles and Francesca Molinari. (2010). "Rounding Probabilistic Expectations in Surveys," *Journal of Business and Economic Statistics*, 28(2): 219-231.

McKenzie, David, John Gibson and Steven Stillman (2007) "A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?" *World Bank Policy Research Working Paper No. 4141.*

Press, W. H.(1992) *Numerical Recipes in C.* Cambridge University Press, Cambridge, UK.

Tversky, Amos and Daniel Kahneman. (1974). "Judgment under uncertainty: Heuristics and biases" *Science*, 185: 1124-1130.

Table 1: Summary Statistics

				Percentiles		
Description variable	**N**	**Mean**	**S. D.**	**10**	**50**	**90**
Age	272	41.42	9.25	30	41	55
Number of children	272	4.79	3.00	0	4	8
Years of Education	271	5.51	2.14	3	5	8
Ability to read a newspaper and write a letter (1=Yes)	272	0.72	0.45	0	1	1
Months as boatowner	272	174.38	93.22	66	159	303
Number of Crew (including self)	272	2.75	1.70	0	3	4

Table 2: Distribution of respondents in each random design

Number of beans	Placement of the pre-determined support expectation question	Placement of the expected maximum question	N
10	First	First	32
10	First	Second	34
10	Second	First	34
10	Second	Second	34
20	First	First	34
20	First	Second	34
20	Second	First	34
20	Second	Second	34
Total			270

Table 3: Comparison of distributions obtained with 10 versus 20 beans

Panel A: Subjective Probability of Going to Town in 2 days

								P-values for testing equality of:			
			Percentiles					Means	Medians	Distribution	
	Mean	S.D.	5	25	50	75	95			K-S	M-W
10 beans	0.456	0.168	0.2	0.3	0.5	0.6	0.7	0.671	0.635	0.139	0.775
20 beans	0.464	0.139	0.25	0.4	0.5	0.55	0.7				

Panel B: Probability allocated to each interval using self-anchored support

			Percentiles					P-values for testing equality of:			
	Mean	S.D.	5	25	50	75	95	Means	Median	Distribution	
Interval 1										K-S	M-W
10 beans	0.343	0.108	0.2	0.3	0.3	0.4	0.5	0.188	0.468	0.040	0.148
20 beans	0.360	0.106	0.2	0.3	0.35	0.45	0.55				
Interval 2											
10 beans	0.339	0.099	0.2	0.3	0.3	0.4	0.5	0.784	0.176	0.515	0.800
20 beans	0.335	0.090	0.2	0.25	0.35	0.4	0.5				
Interval 3											
10 beans	0.191	0.088	0.1	0.1	0.2	0.2	0.3	0.759	0.079	0.047	0.738
20 beans	0.194	0.078	0.1	0.15	0.2	0.25	0.35				
Interval 4											
10 beans	0.127	0.050	0.1	0.1	0.1	0.15	0.2	0.003	0.538	0.011	0.018
20 beans	0.110	0.043	0.05	0.1	0.1	0.15	0.2				

Panel C: Probability allocated to intervals within pre-determined support

			Percentiles					P-values for testing equality of:			
	Mean	S.D.	5	25	50	75	95	Means	Median	Distribution	
First 2 intervals										K-S	M-W
10 beans	0.010	0.030	0	0	0	0	0.1	0.384	0.003	0.089	0.008
20 beans	0.013	0.023	0	0	0	0	0.05				
Intervals 3-10											
10 beans	0.446	0.158	0.2	0.3	0.4	0.6	0.7	0.753	0.107	0.395	0.976
20 beans	0.440	0.146	0.25	0.35	0.4	0.55	0.7				
Intervals 11-18											
10 beans	0.467	0.130	0.2	0.4	0.5	0.6	0.6	0.826	0.230	0.223	0.887
20 beans	0.471	0.126	0.25	0.4	0.5	0.55	0.65				
Last 2 intervals											
10 beans	0.072	0.090	0	0	0	0.1	0.2	0.277	0.088	0.180	0.872
20 beans	0.062	0.069	0	0	0.05	0.1	0.2				
Number of intervals used											
10 beans	9.288	0.648	8	9	9	10	10	0.000	0.000	0.000	0.000
20 beans	13.761	0.935	12	13	14	14	15				

Note: K-S and M-W denote Kolmogorov-Smirnov test and the Mann-Whitney rank-sum test respectively.

Table 4: Comparison of Distributions Elicited under Self-anchored support to those under Pre-determined support

Panel A: Comparing the Distributions of Moments and Percentiles

			Percentiles					P-values for testing equality of:			
	Mean	S.D.	5	25	50	75	95	Means	Median	Distribution	
										K-S	M-W
Mean value of catches											
Self-anchored support	1909.6	488.6	1218.8	1468.1	1860.0	2361.9	2710.0	0.991	0.806	0.799	0.969
Pre-determined support	1908.9	469.1	1257.5	1465.0	1917.5	2247.5	2637.5				
Median value of catches											
Self-anchored support	1749.4	435.5	1100.0	1396.9	1701.4	2071.4	2468.8	0.037	0.014	0.054	0.040
Pre-determined support	1641.9	399.7	1050.0	1312.5	1650.0	1950.0	2366.7				

Panel B: Comparing the amount of probability mass placed in similar intervals

			Percentiles					P-values for testing equality of:			
	Mean	S.D.	5	25	50	75	95	Means	Median	Distribution	
										K-S	M-W
Interval 1											
Self-anchored support	0.351	0.104	0.2	0.3	0.35	0.4	0.55	0.421	0.927	0.844	0.551
Pre-determined support	0.363	0.120	0.2	0.3	0.35	0.4	0.6				
Interval 2											
Self-anchored support	0.336	0.097	0.2	0.25	0.3	0.4	0.5	0.731	0.616	0.994	0.768
Pre-determined support	0.332	0.104	0.2	0.25	0.3	0.4	0.5				
Interval 3											
Self-anchored support	0.199	0.088	0.1	0.15	0.2	0.25	0.4	0.532	0.796	0.933	0.421
Pre-determined support	0.192	0.092	0.05	0.1	0.2	0.25	0.35				
Interval 4											
Self-anchored support	0.114	0.042	0.05	0.1	0.1	0.1	0.2	0.995	0.479	0.539	0.668
Pre-determined support	0.114	0.058	0.05	0.1	0.1	0.15	0.2				

Note: K-S and M-W denote Kolmogorov-Smirnov test and the Mann-Whitney rank-sum test respectively.

Table 5. Percentiles and mean of the subjective minimum and maximum according to the ordering of the questions

	Asked about minimum and maximum first		Asked about pre-determined support distribution first	
	Subjective minimum	Subjective maximum	Subjective minimum	Subjective maximum
5th Percentile	100	2700	100	2600
25th Percentile	200	3500	200	3500
Median	200	4700	200	4500
75th Percentile	300	5500	300	5500
95th Percentile	400	6500	400	7000
Mean	234.31	4558.02	235.74	4558.89
Number of Observations	131	131	135	135

Table 6: Actual distribution compared to Subjective Distribution on Pre-determined support

			Percentiles				
	Mean	S.D.	5	25	50	75	95
First 2 intervals (<300 rupees)							
10 beans	0.010	0.030	0	0	0	0	0.1
20 beans	0.013	0.023	0	0	0	0	0.05
Realizations	0.013	0.031	0	0	0	0	0.08
Intervals 3-10 (300-1500 rupees)							
10 beans	0.446	0.158	0.2	0.3	0.4	0.6	0.7
20 beans	0.440	0.146	0.25	0.35	0.4	0.55	0.7
Realizations	0.506	0.147	0.27	0.41	0.52	0.61	0.68
Intervals 11-18 (1501-4000 rupees)							
10 beans	0.467	0.130	0.2	0.4	0.5	0.6	0.6
20 beans	0.471	0.126	0.25	0.4	0.5	0.55	0.65
Realizations	0.438	0.126	0.24	0.35	0.44	0.52	0.65
Last 2 intervals (>4000 rupees)							
10 beans	0.072	0.090	0	0	0	0.1	0.2
20 beans	0.062	0.069	0	0	0.05	0.1	0.2
Realizations	0.043	0.071	0	0	0	0.08	0.17

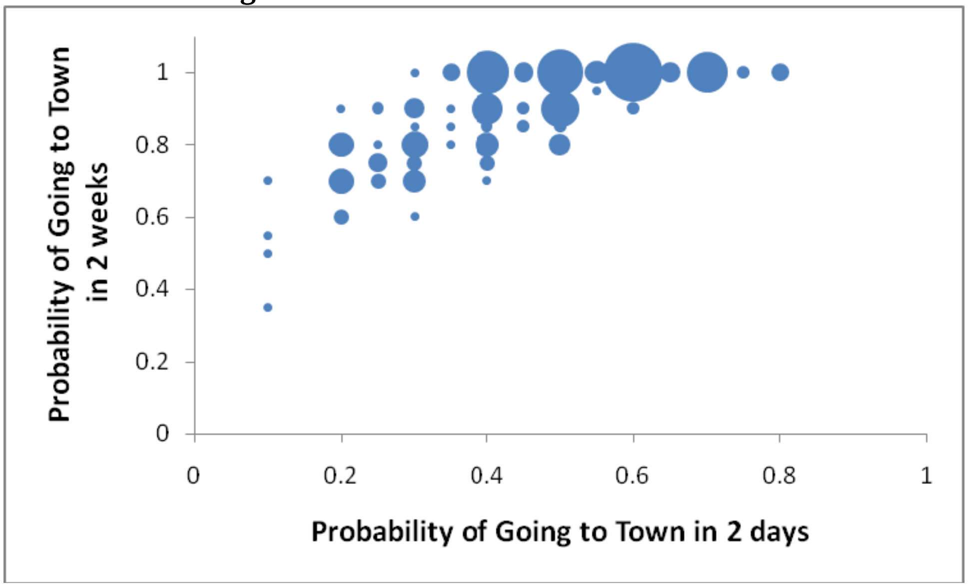
Table 7: Which Choices are more accurate?

	Both Supports				Pre-determined			Self-Reported		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Support is pre-determined (1=Yes)	-0.245***	-0.252***	-0.247***	-0.209***						
	(0.034)	(0.028)	(0.027)	(0.031)						
Number of stones is 20 (1=Yes)	-0.099	-0.121**	-0.130**	-0.092*	-0.136*	-0.162***	-0.173***	-0.060	-0.079	-0.086*
	(0.069)	(0.054)	(0.054)	(0.050)	(0.078)	(0.061)	(0.061)	(0.062)	(0.049)	(0.049)
First distribution elicted (1=Yes)	0.002	0.002	0.002	0.002	0.060	0.083	0.079	-0.057	-0.075	-0.072
	(0.015)	(0.015)	(0.015)	(0.015)	(0.078)	(0.061)	(0.060)	(0.062)	(0.048)	(0.048)
Maximum was elicited first (1=Yes)	-0.072	-0.086*	-0.076	-0.076				-0.071	-0.083*	-0.077
	(0.062)	(0.049)	(0.047)	(0.047)				(0.062)	(0.049)	(0.047)
Pre-determined Support x 20 stones				-0.077***						
				(0.029)						
Age			-0.010	-0.010			-0.008			-0.015
			(0.021)	(0.021)			(0.024)			(0.020)
Age squared			0.000	0.000			0.000			0.000
			(0.000)	(0.000)			(0.000)			(0.000)
Years of Education			-0.106**	-0.106**			-0.123**			-0.083*
			(0.045)	(0.045)			(0.049)			(0.043)
Years of Education squared			0.008**	0.008**			0.010**			0.006
			(0.004)	(0.004)			(0.004)			(0.004)
Has a Fiber Boat (1=Yes)			-0.092	-0.092			-0.144			-0.044
			(0.095)	(0.095)			(0.115)			(0.081)
Has more than one Fiber Boat (1=Yes)			0.232	0.232			0.269			0.185
			(0.154)	(0.154)			(0.165)			(0.145)
Village Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	526	526	524	524	263	263	262	263	263	262
R-squared	0.042	0.413	0.439	0.440	0.014	0.401	0.432	0.012	0.409	0.432

Standard errors clustered at the boat owner level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the logarithm of the absolute value of the area between the cumulative distribution function of actual catches and the elicited distribution. Higher values indicate less accuracy.

Columns 1-4 include the distributions elicited using both pre-determined and self-reported support. Columns 5-7 report only distributions elicited with the pre-determined support while columns 8-10 use distributions elicited with the self-reported support.

Figure 1: Probabilities of Nested Events



Note: Size of bubbles indicates number of data points (ranging from 1, to 44)

Figure 2: Comparison of the Means of the Realized and Subjective Distributions

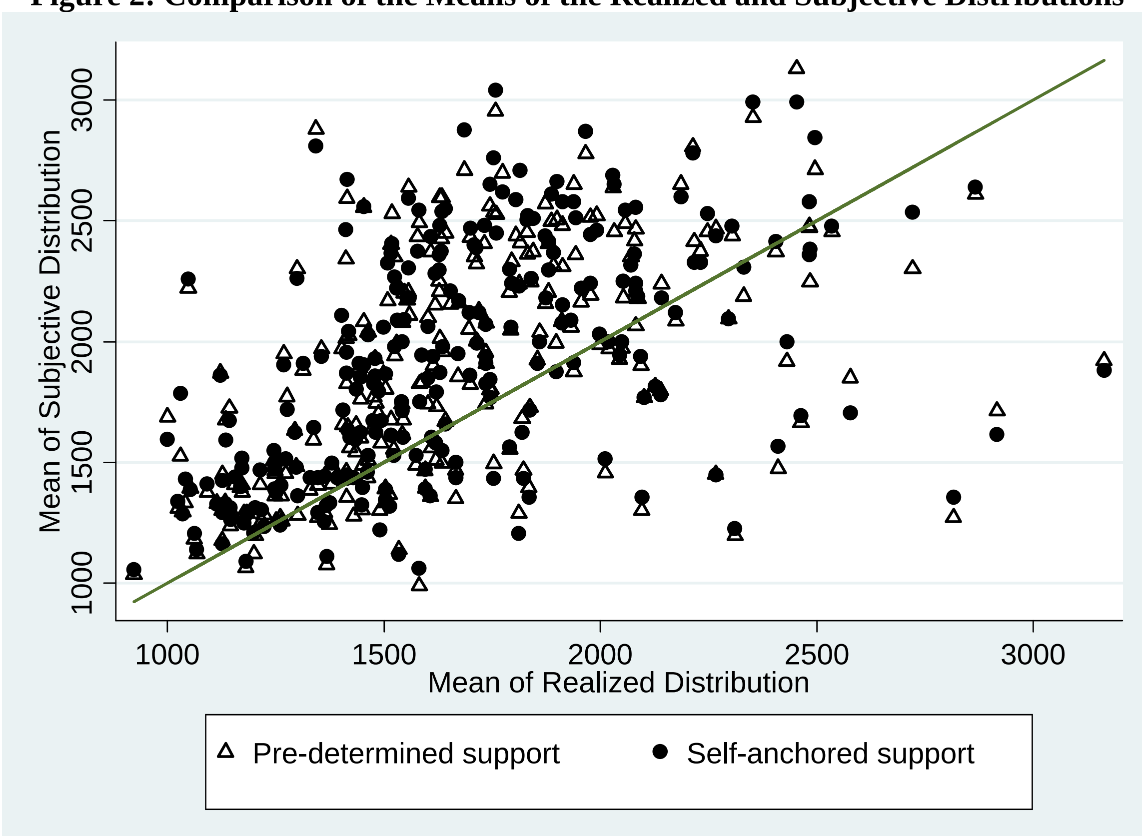


Figure 3: Comparison of the Standard Deviations of the Realized and Subjective Distributions